Genetic Polymorphisms in the *IGFBP3* Gene: Association with Breast Cancer Risk and Blood IGFBP-3 Protein Levels among Chinese Women

Zefang Ren,[1] Qiuyin Cai,[1] Xiao-Ou Shu,[1] Hui Cai,[1] Chun Li,[2] Herbert Yu,[3] Yu-Tang Gao,[4] and Wei Zheng[1]

[1]Department of Medicine and Vanderbilt-Ingram Cancer Center and [2]Department of Molecular Physiology and Biophysics, Vanderbilt University, Nashville, Tennessee; [3]Department of Epidemiology and Public Health, Yale Cancer Center, Yale University School of Medicine, New Haven, Connecticut; and [4]Department of Epidemiology, Shanghai Cancer Institute, Shanghai, China

Abstract

Cumulative evidence suggests that insulin-like growth factors (IGF) play an important role in the etiology of breast cancer. The IGF binding proteins regulate the action of IGFs, and >90% of circulating IGFs are bound to IGFBP-3. We evaluated the associations of five (*A−202C*, *G227C*, *C3804G*, *5606InsA*, and *C5827T*) genetic polymorphisms in the *IGFBP3* gene with breast cancer risk and the blood IGFBP-3 protein level in a population-based, case-control study conducted among Chinese women in Shanghai. Genomic DNA samples from 1,193 incident breast cancer patients and 1,310 community controls were genotyped for *IGFBP3* polymorphisms. Blood IGFBP-3 levels were determined for 390 controls. A 30% to 60% elevated risk of breast cancer was found to be associated with homozygosity for the variant allele in polymorphisms *A−202C*, *G227C*, *5606InsA*, and *C5827T*. Carrying the variant allele in *C3804G* was also associated with an increased risk. About 13.5% of cases and 9.7% of controls had one or more of the above risk genotypes, resulting in odds ratio [OR; 95% confidence interval (95% CI)] of 1.4 (1.0-1.9). The ORs (95% CIs) were 1.3 (1.0-1.8) and 1.7 (1.1-2.5) for women with one to two and three to five risk genotypes, respectively (*P* for trend < 0.01). Four common haplotypes for the *IGFBP3* gene were identified. Compared with the haplotype containing only the wild-type allele in the five loci, the haplotype with the variant allele in all sites was associated with an elevated risk of breast cancer (OR 1.4, 95% CI 1.0-1.9), particularly among younger women (OR 2.3, 95% CI 1.3-3.9). With the exception of *C3804G*, in which no homozygote was identified, the level of circulating IGFBP-3 was reduced in a dose-response manner with an increasing number of variant alleles in each of the other four polymorphic sites (*P* for trend < 0.05). These results indicated that *IGFBP3* polymorphisms may be associated with the level of blood IGFBP-3 protein and an increased risk of breast cancer. (Cancer Epidemiol Biomarkers Prev 2004;13(8):1290–5)

Introduction

The insulin-like growth factor (IGF) family includes the polypeptide ligands IGF-I and IGF-II, the IGF receptors, and six binding IGF proteins (i.e., IGFBP-1 to IGFBP-6; refs. 1-3). A large number of *in vitro* studies have shown that IGFs are strong mitogens for a variety of cancer cells including many breast cancer cell lines (1-3). Besides mitogenic stimulation, IGFs also inhibit apoptosis (1-3). The combination of these mitogenic and antiapoptotic effects has a profound impact on tumor growth (1-3). IGF-I and IGF-II are present in the circulation where the majority of them (>90%) are bound to IGFBP-3 and become less available to the target tissues (1-3). On the other hand, the binding of IGF-I/IGF-II to IGFBP-3 prevents these growth factors from degradation and, in some situations, may enhance the action of IGFs (1-3). IGFBP-3 also directly binds to its putative cell surface receptor and induces apoptosis independent of IGF-I/IGF-II (1-4). It has been reported that IGFBP-3 can mediate the growth inhibitory action of transforming growth factor-1 and induce apoptosis through the tumor suppressor gene *p53* in breast cancer cells (2, 3).

Although IGF-I levels in the blood have been fairly consistently shown to be positively associated with premenopausal breast cancer risk in previous epidemiologic studies (1, 5-8), the reports on IGFBP-3 have been conflicting (1, 5-13). Both positive and inverse associations of breast cancer with this molecule have been reported. The levels of IGFBP-3 can be affected by both lifestyle and genetic factors (14-16). Recently, a polymorphism ($A \rightarrow C$ transversion) in the promoter region of the *IGFBP3* gene was reported. It was found that the variant genotype was associated with a significantly decreased level of IGFBP-3 in both *in vitro* and *in vivo* studies (15, 16). Several other common polymorphisms have also been reported in the *IGFBP3* gene (17-19), including a $G \rightarrow C$ transversion in exon 1 (Gly → Ala substitution at codon 32; refs. 17, 18), a $C \rightarrow G$ transversion located only 17 nucleotides upstream of

Received 11/21/03; revised 3/4/04; accepted 3/11/04.

Grant support: USPHS grants RO1CA64277 and RO1CA90899 from the National Cancer Institute.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Note: Zefang Ren is a visiting postdoctoral fellow from the School of Public Health, Sun Yat-Sen University, Guang Zhou, Guandong, China.

Requests for reprints: Wei Zheng, Center for Health Services Research, Medical Center East, Suite 6000, Nashville, TN 37232-8300. Phone: 615-936-0682; Fax: 615-936-1269. E-mail: wei.zheng@vanderbilt.edu

exon 2 (17, 18), and an $A \rightarrow C$ transversion at nucleotide 7,580 in intron 3 (19). Recently, we identified two additional polymorphisms in intron 3 of the *IGFBP3* gene and evaluated, in this study, all these polymorphisms in relation to breast cancer risk.

Methods

Subject Recruitment and Data/Specimen Collection. Included in this study were subjects recruited during 1996 to 1998 in the Shanghai Breast Cancer Study. Detailed study methods have been published elsewhere (20). Briefly, this study included 1,459 incident breast cancer cases diagnosed at an age between 25 and 64 years and 1,556 age frequency-matched community controls. Cancer cases were identified through the population-based Shanghai Cancer Registry, which records virtually all cancer cases diagnosed among residents of urban Shanghai. A total of 1,602 eligible breast cancer cases were identified during the study period, of which 1,459 (91.1%) cases completed in-person interviews. Cancer diagnoses for all patients were reviewed and confirmed by two senior pathologists. Controls were randomly selected from the general population in Shanghai using the Shanghai Resident Registry, a population registry containing demographic information for all residents of urban Shanghai, and were frequency matched on age (5-year intervals) to the expected age distribution of the case subjects in a 1:1 ratio. The inclusion criteria for controls were identical to those of the cases with the exception of a breast cancer diagnosis. Of the 1,724 eligible women, 1,556 (90.3%) completed in-person interviews. The major reason for nonparticipation was refusal, accounting for 6.8% (n = 109) for cases and 9.6% (n = 166) of controls identified for the study.

A structured questionnaire was used to elicit detailed information on demographic factors, menstrual and reproductive history, hormone use, dietary habits, prior disease history, physical activity, tobacco and alcohol use, weight, and family history of cancer. All participants were measured for their current weight and circumference of waist and hips. Blood samples (10 mL from each woman) were obtained from 1,193 (82%) cases and 1,310 (84%) controls who completed the in-person interviews. These samples were processed on the same day, typically within 6 hours of sample collection, and were stored at −70°C until relevant bioassays.

Laboratory Protocols. The laboratory work was conducted during 2002 at the Molecular Epidemiology Laboratory of the Vanderbilt-Ingram Cancer Center (Nashville, TN). Genomic DNA was extracted from buffy coats with the use of the Puregene DNA Purification Kit (Gentra Systems, Minneapolis, MN) following the manufacturer's protocol. DNA concentration was measured using PicoGreen dsDNA Quantitation Kit (Molecular Probes, Eugene, OR). Genotyping was performed using PCR-based RFLP assays. The RFLP method failed to reveal the polymorphism ($A \rightarrow C$) at nucleotide 7,580 reported by Sun et al. (19). To confirm this finding, we sequenced 10 study samples [from nucleotides 7,343 to 7,953 based on the Genbank accession no. M35878 sequence or from nucleotides 5,438 to 6,048 relative to transcription start site (21)]. The primers used for PCR amplification were 5′-CTCCGAGTCACTGGCATTTC-3′ and 5′-ACCAGCCCTTGTAGAACCTC-3′ (19). The PCR reactions were performed in a 20 μL mixture containing 5 ng template DNA, 0.5 unit Hotstar Taq DNA polymerase (Qiagen, Valencia, CA), 1× Qiagen PCR buffer, 1.5 mmol/L $MgCl_2$, 0.2 mmol/L each of deoxynucleotide triphosphates, and 0.5 μmol/L each primer. After denaturation at 95°C for 15 minutes, the PCR was performed in 35 cycles of 94°C for 45 seconds, 55°C for 45 seconds, and 72°C for 45 seconds. At the end, the reactions were extended for 7 minutes at 72°C. The PCR products were purified with a Qiagen PCR purification kit and directly sequenced with the Big Dye Terminator sequencing chemistry in ABI PRISM 3700 DNA analyzer (Applied Biosystems, Foster City, CA) for both DNA strands of the 10 study samples. Results from direct DNA sequencing confirm that the *A7580C* polymorphism does not exist in our study population. Two novel polymorphisms, however, were found when compared with the original *IGFBP3* gene sequence in Genbank: an insertion of adenine at nucleotide 5,606 (*5606InsA*) and a transition from $C \rightarrow T$ at nucleotide 5,827 (*C5827T*) relative to transcription start site (21). These two polymorphisms, along with the three previously reported polymorphisms (*A−202C*, *G227C*, and *C3804G*) that were confirmed in our study population, were included in this study.

Table 1 shows PCR primers, restriction enzymes, and length of the resulting fragments in each genotype. The PCR conditions for genotyping were similar to that described above for DNA sequencing, except for a different annealing temperature and the addition of Q-solution

Table 1. Summary of methods for the five polymorphisms evaluated in the study

Locations*		Sequence change	PCR primers	Annealing temperature (°C)	Enzyme	Alleles	Restriction fragments (bp)
A	B						
−202	1,704	$A \rightarrow C$	F: 5′-GAATGCGGAGCGCTGTATG-3′ R: 5′-TGTGGAATCCAGGCAGGAAG-3′	59	*Fsp*I	*A* *C*	483 258,225
227	2,132	$G \rightarrow C$ (*G32A*)	F: 5′-GCCTGGATTCCACAGCTTC-3′ R: 5′-TGACGGCACTAGCGTTGAC-3′	55	*Ava*I	*G* *C*	271,179 450
3,804	5,709	$C \rightarrow G$	F: 5′-ACTGAAGCGTTTGTCTGCCT-3′ R: 5′-AACTTTGTAGCGCTGGCTGT-3′	52	*Bsm*AI	*C* *G*	264,98,77 341,98
5,606	7,511	Insertion A	F: 5′-CTCCGAGTCACTGGCATTTC-3′ R: 5′-ACCAGCCCTTGTAGAACCTC-3′	55	*Nde*I	No insertion A Insertion A	611 443,168
5,827	7,732	$C \rightarrow T$	Identical to those for 7,511 insertion A	55	*Bts*I	*C* *T*	397,214 661

*A, location relative to the transcription start site described by Cubbage et al. (21); B, location based on the sequence from Genbank accession no. M35878.

(Qiagen) for *A−202C* and *G227C* that are in the GC-rich region. Ten microliters of each PCR product were digested with 5 units of the corresponding restriction enzyme (NEB, Beverly, MA) following the manufacturer's instructions in a final reaction volume of 20 μL. The restriction fragments were separated by gel electrophoresis on a 1.5% agarose gel containing ethidium bromide; the gel was visualized and pictured on a UV transilluminator. Because of insufficient genomic DNA or laboratory failure, genotyping data were unavailable for 6.6% of study samples.

Plasma concentrations of IGFBP-3 were measured in 390 control women who were included in an ancillary study conducted as part of the Shanghai Breast Cancer Study (6). A commercially available ELISA kit (DSL Inc., Webster, TX) was used in the assay as has been used in the majority of previous epidemiologic studies (1, 5-8). The coefficients of variation for intraassay and interassay precisions were 0.5% to 1.9% and 1.8% to 3.9%, respectively.

Statistical Analysis. χ^2 statistics were used to evaluate case-control differences in the distribution of *IGFBP3* genotypes. Odds ratios (OR) and 95% confidence intervals (CI) were used to measure the strength of the association between *IGFBP3* polymorphisms and breast cancer risk after adjusting for age and other potential confounding factors in logistic regression models (22). Stratified analyses based on age (<45 and ≥45 years) were performed to examine whether the association was stronger among younger women. Tests for Hardy-Weinberg equilibrium were conducted by comparing observed and expected frequencies of genotypes using the χ^2 test. The linkage disequilibrium level D' was computed for every pair of polymorphisms using the software GOLD (23). Assuming Hardy-Weinberg equilibrium, haplotype frequencies for case and control groups were estimated (24), and the ORs for common haplotypes were calculated in comparison with the haplotype carrying only the most frequent alleles in all five loci evaluated in this study. ANOVA was used to compare plasma IGFBP-3 levels by genotypes. $P < 0.05$ (two-sided probability) was interpreted as statistically significant.

Results

Results on the demographic variables and known breast cancer risk factors in the study were published previously (20). Breast cancer patients and controls were comparable in age and education level. With the exception of a family history of breast cancer, statistically significant associations were observed for all major risk factors for breast cancer, including early age at menarche, late age at menopause, late age at first live birth, physical inactivity, a prior history of breast fibroadenoma, high body mass index, and high waist-to-hip ratio. More cases than controls had a family history of breast cancer, although the differences were not statistically significant, perhaps due to a few subjects having a positive family history of breast cancer.

The associations between *IGFBP3* genotypes and breast cancer risk are presented in Table 2. No subjects were homozygous for the *G* allele at nucleotide 3,804, and an elevated OR of borderline significance was observed among women heterozygous for the *G* allele. For each of the other four polymorphisms, elevated risks of breast cancer were observed among women who were homozygous for the variant allele, and three of the four age-adjusted ORs were statistically significant. The ORs were essentially unchanged after adjusting for the nongenetic breast cancer risk factors that are listed in the footnote of Table 2. Only a small percentage of women had used hormone replacement therapy (2.5%) or had a family history of breast cancer (2.4%). Adjusting for these variables did not appreciably affect the risk estimate. Additional analyses were performed to evaluate the association of breast cancer risk with having more than one risk genotype. Compared with women who did not have any risk genotype, those with one to two and three to five risk genotypes had elevated risks of 1.3 and 1.7, respectively, and the trend for increases in risk was statistically significant ($P < 0.01$). When stratified by age, the positive association appeared to be stronger among younger women (OR 1.7, 95% CI 1.1-2.5) than older women (OR 1.2, 95% CI 0.8-1.7).

The degree of linkage disequilibrium was evaluated for every pair of polymorphism. The D' values ranged from 0.88 to 1.00, indicating that the five polymorphisms under investigation are in close linkage disequilibrium. Because the interaction of multiple polymorphisms within a haplotype could potentially affect biological phenotypes, the frequencies of extended haplotypes were estimated, and their associations with breast cancer risk were evaluated (Table 3). Four common haplotypes (frequency of >1%) were found to account for 96.5% of chromosomes in the study population. When compared with the haplotype containing only the most frequent allele in all five polymorphic sites, a 40% elevated risk of breast cancer was found to be associated with the haplotype containing all variant alleles. The positive association was found primarily among younger women, with OR (95% CI) of 2.3 (1.3-3.9). No elevated risk of breast cancer, however, was found to be associated with other haplotypes containing one to four variant alleles.

To assess potential joint effects of *IGFBP3* polymorphisms and factors related to endogenous estrogen exposure, subjects were classified based on the joint distribution of *IGFBP3* polymorphisms and estrogen-related factors, including waist-to-hip ratio, body mass index, age at first live birth, age at menarche, age at menopause, and regular physical activity. The risks of breast cancer associated with *IGFBP3* polymorphisms were elevated in virtually all strata defined by estrogen-related factors (data not shown in table). None of the tests based on multiplicative models, however, were statistically significant.

Figure 1 shows plasma IGFBP-3 protein levels among controls by *IGFBP3* genotypes. Again, no subjects were homozygous for the variant *G* allele in *C3804G*. The age-adjusted plasma IGFBP-3 level was lower in heterozygotes than homozygotes of the *C* allele (wild-type). For each of the other four polymorphisms, women homozygous for the wild-type allele had the highest mean of blood IGFBP-3, and the level declined in a stepwise manner with the presence of one and two copies of the variant alleles (P for trend < 0.05). There were statistically significant differences between wild-type and variant homozygous genotypes in all four polymorphisms but none between heterozygotes and homozygotes. When all polymorphisms were considered, the mean plasma IGFBP-3

Table 2. *IGFBP3* allele and genotype frequencies and ORs (95% CIs) for breast cancer risk, the Shanghai Breast Cancer Study

Polymorphism	Genotypes	Case patients	Control subjects	OR* (95% CI)	OR† (95% CI)
A−202C	*AA*	641	708	1.0 (reference)	1.0 (reference)
	AC	405	455	1.0 (0.8-1.2)	1.0 (0.8-1.2)
	CC	63	44	1.6 (1.1-2.4)	1.6 (1.0-2.3)
G227C	*GG*	653	739	1.0 (reference)	1.0 (reference)
	GC	391	430	1.0 (0.9-1.2)	1.0 (0.9-1.2)
	CC	63	42	1.7 (1.1-2.5)	1.6 (1.1-2.5)
C3804G	*CC*	1,051	1,167	1.0 (reference)	1.0 (reference)
	GC	84	69	1.4 (1.0-1.9)	1.3 (0.9-1.9)
	GG	0	0		
5606InsA	*NN*‡	596	655	1.0 (reference)	1.0 (reference)
	NA	447	502	1.0 (0.8-1.2)	1.0 (0.8-1.2)
	AA	78	58	1.4 (1.0-2.1)	1.4 (1.0-2.1)
C5827T	*CC*	611	669	1.0 (reference)	1.0 (reference)
	CT	446	488	1.0 (0.9-1.2)	1.0 (0.8-1.2)
	TT	72	58	1.3 (0.9-1.9)	1.3 (0.9-1.9)
All women					
Presence of any risk genotypes§					
No		941	1,061	1.0 (reference)	1.0 (reference)
Yes		148	114	1.4 (1.1-1.9)	1.4 (1.0-1.8)
No. of risk genotypes					
1-2		90	75	1.3 (1.0-1.8)	1.3 (0.9-1.8)
3-5		58	39	1.7 (1.1-2.5)	1.6 (1.1-2.5)
P for trend				0.01	0.01
Stratified analyses by age					
Women <45 y					
Presence of any risk genotypes					
No		376	442	1.0 (reference)	1.0 (reference)
Yes		66	46	1.6 (1.1-2.5)	1.7 (1.1-2.5)
Women ≥45 y					
Presence of any risk genotypes					
No		565	619	1.0 (reference)	1.0 (reference)
Yes		82	68	1.3 (0.9-1.9)	1.2 (0.9-1.8)

*OR adjusted for age.
†OR adjusted for age, physical activity, waist-to-hip ratio, body mass index, age at menarche, nulliparity, age at first live birth, and menopausal status.
‡*N*, no insertion.
§Risk genotypes include *CC* at *A−202C*, *CC* at *G227C*, *CG* at *C3804G*, *AA* at *5606InsA*, and *TT* at *C5827T*.

protein was the highest in subjects homozygous for the wild-type allele in all five loci ($n = 235$, $\bar{x} \pm SD = 4{,}153 \pm 1{,}658$) followed by heterozygotes in all but *C3804G* ($n = 105$, $\bar{x} \pm SD = 3{,}774 \pm 1{,}195$) and those heterozygous for the variant allele in *C3804G* or homozygous for the variant allele in any one of the other four sites ($n = 32$, $\bar{x} \pm SD = 3{,}417 \pm 1{,}022$). The trend test for this association was highly significant ($P = 0.002$; data not shown in Fig. 1). The level of blood IGFBP-3 protein appeared to reduce further with an increased number of risk genotypes (mean = 3,465 for one to two risk genotype groups and mean = 3,326 for three to five risk genotype groups). The sample size, however, was small, and the difference between these two groups was not statistically significant.

Table 3. Analysis of common haplotypes in the *IGFBP3* gene: frequency estimates and ORs for their associations with breast cancer risk, the Shanghai Breast Cancer Study

Haplotypes*	All subjects			Age < 45 y			Age ≥ 45 y		
	Case patients (%)	Control subjects (%)	OR (95% CI)	Case patients (%)	Control subjects (%)	OR (95% CI)	Case patients (%)	Control subjects (%)	OR (95% CI)
A, G, C, N, C	71.70	73.70	1.0 (reference)	72.13	74.63	1.0 (reference)	71.41	73.03	1.0 (reference)
A, G, C, A, T	2.80	3.19	0.9 (0.6-1.3)	2.31	3.13	0.7 (0.4-1.3)	3.15	3.23	1.0 (0.7-1.6)
C, C, C, A, T	17.92	17.46	1.0 (0.9-1.2)	17.16	16.77	1.1 (0.8-1.4)	18.43	17.96	1.1 (0.9-1.3)
C, C, G, A, T	3.48	2.64	1.4 (1.0-1.9)	4.52	2.05	2.3 (1.3-3.9)	2.78	3.06	0.9 (0.6-1.5)
No. of chromosomes	2,178	2,350		884	976		1,294	1,374	

*From left, the polymorphic sites are *A−202C*, *G227C*, *C3804G*, *5606InsA*, and *C5827T*. *N*, no insertion at nucleotide 5606.

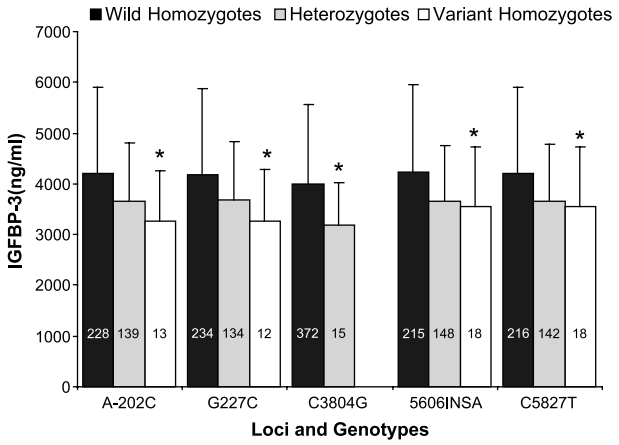

Figure 1. Blood levels ($\bar{x} \pm$ SD) of IGFBP-3 in control subjects by genotypes. *Numbers in bars,* numbers of control subjects. (*, $P < 0.05$, compared with wild-type homozygotes. Tests for linear trend of blood IGFBP-3 levels across genotypes were statistically significant at $P < 0.05$ for *A−202C*, *G227C*, *5606InsA*, and *C5827T*).

Discussion

Unlike most Mendelian disorders that are caused by the mutation of a single gene, breast cancer is a complex disease, resulting from the insult of multiple genetic and environmental factors (25-27). Germ line mutation of the *BRCA1* and *BRCA2* genes accounts for <10% of breast cancer cases in the general population (26, 27). Over the past 5 years, a great deal of research has been focused on identifying genetic factors that are responsible for breast cancer, particularly the genes that are involved in estrogen and carcinogen metabolism (26). Because of their high frequencies in the general population and their potential interactions with environmental factors, these genes may be responsible for a high proportion of breast cancer cases (26, 27). In this large, population-based, case-control study, we found significant associations of genetic polymorphisms in the *IGFBP3* gene with the risk of breast cancer and a correlation between *IGFBP3* genotype and phenotype. These findings suggest that *IGFBP3* polymorphisms may be an important genetic factor for breast cancer susceptibility.

Several epidemiologic studies have been conducted to evaluate the association of blood IGFBP-3 levels with breast cancer risk along with other IGF-related molecules. The results, however, have been conflicting. Several case-control studies reported an inverse association (9, 10), whereas others reported a positive association (6, 11, 13). Results from prospective studies have also been inconsistent. Both positive (7) and inverse (5, 7) associations have been reported. Two prospective cohort studies showed no association of IGFBP-3 levels with breast cancer risk (8, 12). We recently reported that blood IGFBP-3 level was positively associated with the risk of breast cancer in a subset of participants from the Shanghai Breast Cancer Study (6). This result appears contradictory to the findings from the current study, in which the risk of breast cancer was associated with the genotypes related to reduced blood IGFBP-3 levels. The conflicting results from epidemiologic studies on blood IGFBP-3 levels are perhaps not unexpected, given the dual roles of IGFBP-3 protein in regulating the actions of IGF-I. The net effect of IGFBP-3 on IGF-I is heavily determined by the activity of IGFBP protease, particularly at the target tissue. Blood level of IGFBP-3 may not reflect the level and action of this protein in the target tissues, and most epidemiologic studies have no access to normal target tissue samples to evaluate the association of this protein with cancer risk. Therefore, genetic polymorphisms in the *IGFBP3* gene could be a better indicator of the level of IGFBP-3 protein in the target tissues than that in the circulation. Our findings for a reduced blood IGFBP-3 level and an elevated risk of breast cancer among subjects carrying the variant alleles of this gene are consistent with the evidence from most *in vitro* experiments and some epidemiologic studies. Because of the limitations of using blood IGFBP-3 levels as the surrogate measure of the level of this protein in the target tissue, caution should be exercised in interpreting findings based on blood IGFBP-3 measurement.

It was recently reported that a transversion from *A* to *C* at nucleotide −202 in the promoter region of the *IGFBP3* gene significantly decreased the expression of this gene in an *in vitro* study (16). This polymorphism was also found to be associated with a reduced level of blood IGFBP-3 protein in three human studies (15, 16, 28). In agreement with these reports, we observed in our study a gene-dose association between blood IGFBP-3 levels and the number of variant *C* alleles in this polymorphism. Furthermore, we have shown in this study that women homozygous for the *C* allele were at an elevated risk of breast cancer, whereas heterozygosity for the *C* allele was unrelated to risk. The reduction in the levels of circulating IGFBP-3 protein, however, was less evident in women heterozygous for the *C* allele than those homozygous for this allele.

Similar positive associations were observed in our study for the other four polymorphisms, although there are currently no *in vitro* data available to support the functionality of these polymorphisms. The $G \rightarrow C$ transversion at nucleotide 227 results in a change of Gly → Ala in codon 32 (exon 1). The other three polymorphisms were identified in introns, which could affect the splicing site on mRNA after transcription. In particular, the polymorphism ($C \rightarrow G$) at nucleotide 3,804 is located only 17 nucleotides upstream of exon 2, within the branch site that is important in RNA splicing (25). If more than one of the polymorphisms evaluated in the study were functionally important, we would expect that the risk of breast cancer might be increased with the number of the risk genotypes. In agreement with this notion, we did observe an increase in breast cancer risk with the number of risk genotypes. In the analysis of extended haplotypes of this gene, we found that the risk of breast cancer was only associated with the haplotype that carries the variant allele in all five polymorphic sites, suggesting that more than one of the polymorphisms may be functionally significant.

The frequencies of variant alleles in our study differ from those reported previously in other populations. For the *A−202C* polymorphism, the frequencies for the *C* allele were 0.40 and 0.53 in Caucasian men and women, respectively (15, 16). In our study, however, the frequency of this allele was only 0.23. In a small study of

23 Caucasians, Eggermann et al. found the variant allele frequencies were 0.59 and 0.22 in *G227C* and *C3804G*, respectively (13), again much higher than Chinese women whose frequencies were 0.21 and 0.03, respectively, for these two polymorphisms. These data are consistent with the lower incidence of breast cancer observed among Chinese women when compared with their American counterparts (29). Very recently, Schernhammer et al. (28) reported results from a case-control study evaluating the association between *A-202* polymorphism and breast cancer risk. Although no overall association was found, the *AA* genotype was associated with ~30% reduced risk of breast cancer among premenopausal women <50 years old. However, the sample size for this group of women (70 cases and 70 controls) was small, and the OR was not statistically significant. Nevertheless, the result from this subgroup of women is consistent with the role of IGFBP-3 in breast tumorigenesis and in agreement with the findings from our study, which was conducted in a relatively younger population with a mean age of 47.14 years in controls and 47.65 years in cases.

Noteworthy strengths of this study are its large sample size, population-based design, and a very high response rate, which minimizes potential selection bias. The consistency of the findings for both *IGFBP3* phenotype and haplotype with breast cancer risk, as well as in stratified analyses, strongly indicates that the associations identified in this study are unlikely due to a type I error, a major concern in many association studies of genetic polymorphisms with a small sample size. Over 98% of study subjects belong to a single ethnic group (Han Chinese); thus, possible confounding effects due to population stratification are not a major concern in this study. Extensive survey data were collected in this study, and potential confounding factors were adjusted for in the analysis.

In summary, in this large, carefully conducted, population-based, case-control study, we have shown that *IGFBP3* polymorphisms were associated with the level of circulating IGFBP-3 protein and the risk of breast cancer. These findings are consistent with the cancer inhibitory effect of IGFBP-3 demonstrated in most *in vitro* experiments and some epidemiologic studies showing an inverse association of this protein with breast cancer risk. Our findings are new and may have significant public health implications for identifying high-risk women for the prevention of breast cancer, the most common malignancy in many parts of the world.

Acknowledgments

We thank Drs. Qi Dai, Fan Jin, and Jia-Rong Cheng for contributions in coordinating data and specimen collection in Shanghai, Bethanie Hull for technical assistance in the preparation of this article, and all study participants and research staff of the Shanghai Breast Cancer Study for support.

References

1. Yu H, Rohan T. Role of the insulin-like growth factor family in cancer development and progression. J Natl Cancer Inst 2000;92:1472-89.
2. Schedlich LJ, Graham LD. Role of insulin-like growth factor binding protein-3 in breast cancer cell growth. Microsc Res Tech 2002;59:12-22.
3. Marshman E, Streuli CH. Insulin-like growth factors and insulin-like growth factor binding proteins in mammary gland function. Breast Cancer Res 2002;4:231-9.
4. Hong J, Zhang G, Dong F, Rechler MM. Insulin-like growth factor (IGF)-binding protein-3 mutants that do not bind IGF-I or IGF-II stimulate apoptosis in human prostate cancer cells. J Biol Chem 2002;277:10489-97.
5. Hankinson SE, Willett WC, Colditz GA, et al. Circulating concentrations of insulin-like growth factor-1 and risk of breast cancer. Lancet 1998;351:1393-6.
6. Yu H, Jin F, Shu XO, et al. Insulin-like growth factors and breast cancer risk in Chinese women. Cancer Epidemiol Biomarkers & Prev 2002;11:705-12.
7. Krajcik RA, Borofsky ND, Massardo S, Orentreich N. Insulin-like growth factor I (IGF-I), IGF-binding proteins, and breast cancer. Cancer Epidemiol Biomarkers & Prev 2002;11:1566-73.
8. Toniolo P, Bruning PF, Akhmedkhanov A, et al. Serum insulin-like growth factor-I and breast cancer. Int J Cancer 2000;88:828-32.
9. Bruning PF, Van Doorn J, Bonfrer JM, et al. Insulin-like growth-factor binding protein 3 is decreased in early-stage operable pre-menopausal breast cancer. Int J Cancer 1995;62:266-70.
10. Bohlke K, Cramer DW, Trichopoulos D, Mantzoros CS. Insulin-like growth factor-1 in relation to premenopausal ductal carcinoma *in situ* of the breast. Epidemiology 1998;9:570-3.
11. Del Giudice ME, Fantus IG, Ezzat S, McKeown-Eyssen G, Page D, Goodwin PJ. Insulin and related factors in premenopausal breast cancer risk. Breast Cancer Res Treat 1998;47:111-20.
12. Kaaks R, Lundin E, Rinaldi S, et al. Prospective study of IGF-I, IGF-binding proteins, and breast cancer risk, in northern and southern Sweden. Cancer Causes & Control 2002;13:307-16.
13. Muti P, Quattrin T, Grant BJ, et al. Fasting glucose is a risk factor for breast cancer: a prospective study. Cancer Epidemiol Biomarkers & Prev 2002;11:1361-8.
14. Holmes MD, Pollak MN, Hankinson SE. Lifestyle correlates of plasma insulin-like growth factor I and insulin-like growth factor binding protein 3 concentrations. Cancer Epidemiol Biomarkers & Prev 2002;11:862-7.
15. Jernstrom H, Deal C, Wilkin F, et al. Genetic and nongenetic factors associated with variation of plasma levels of insulin-like growth factor-I and insulin-like growth factor-binding protein-3 in healthy premenopausal women. Cancer Epidemiol Biomarkers & Prev 2001;10:377-84.
16. Deal C, Ma J, Wilkin F, et al. Novel promoter polymorphism in insulin-like growth factor-binding protein-3: correlation with serum levels and interaction with known regulators. J Clin Endocrinol Metab 2001;86:1274-80.
17. Zou T, Fleisher AS, Kong D, et al. Sequence alterations of insulin-like growth factor binding protein 3 in neoplastic and normal gastrointestinal tissues. Cancer Res 1998;58:4802-4.
18. Eggermann K, Wollmann HA, Tomiuk J, Ranke MB, Kaiser P, Eggermann T. Screening for mutations in the promoter and the coding region of the IGFBP1 and IGFBP3 genes in Silver-Russell syndrome patients. Hum Hered 1999;49:123-8.
19. Sun G, Chagnon M, Bouchard C. A common polymorphism in the human insulin-like growth factor binding protein 3 gene. Mol Cell Probes 2000;14:55-6.
20. Gao YT, Shu XO, Dai Q, et al. Association of menstrual and reproductive factors with breast cancer risk: results from the Shanghai Breast Cancer Study. Int J Cancer 2000;87:295-300.
21. Cubbage ML, Suwanichkul A, Powell DR. Insulin-like growth factor binding protein-3: organization of the human chromosomal gene and demonstration of promoter activity. J Biol Chem 1990;265:12642-9.
22. Breslow NE, Day NE. Statistical methods in cancer research. Vol. 1. The analysis of case-control studies (IARC Scientific Publication No. 32). Lyon (France): IARC; 1980. p. 5-338.
23. Abecasis GR, Cookson WO. GOLD—graphical overview of linkage disequilibrium. Bioinformatics 2000;16:182-3.
24. Excoffier L, Slatkin M. Maximum-likelihood estimation of molecular haplotype frequencies in a diploid population. Mol Biol Evol 1995;12:921-7.
25. Strachan T, Read AP. DNA structure and gene expression. In: Kingston F, editor. Human molecular genetics. 2nd ed. New York: BIOS Scientific Publishers Ltd; 1999. p. 14-5.
26. de Jong MM, Nolte IM, te Meerman GJ, et al. Genes other than BRCA1 and BRCA2 involved in breast cancer susceptibility. J Med Genet 2002;39:225-42.
27. Rebbeck TR, Couch FJ, Kant J, et al. Genetic heterogeneity in hereditary breast cancer: role of BRCA1 and BRCA2. Am J Hum Genet 1996;59:547-53.
28. Schernhammer ES, Hankinson SE, Hunter DJ, Blouin MJ, Pollak MN. Polymorphic variation at the −202 locus in IGFBP3: influence on serum levels of insulin-like growth factors, interaction with plasma retinol and vitamin D and breast cancer risk. Int J Cancer 2003;107:60-4.
29. Lacey JV Jr, Devesa SS, Brinton LA. Recent trends in breast cancer incidence and mortality. Environ Mol Mutagen 2002;39:82-8.